

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

**EP 0 824 246 A2**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
18.02.1998 Bulletin 1998/08

(51) Int Cl.<sup>6</sup>: **G06T 7/00**

(21) Application number: **97305609.6**

(22) Date of filing: **25.07.1997**

(84) Designated Contracting States:  
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE**  
Designated Extension States:  
**AL LT LV RO SI**

- Rao, Ramana L.  
Los Alamos, New Mexico 87544 (US)
- Venable, Dennis L.  
Marion NY 14505 (US)
- Eschbach, Reiner  
Webster NY 14580 (US)

(30) Priority: **06.08.1996 US 692559**

(71) Applicant: **XEROX CORPORATION**  
**Rochester New York 14644 (US)**

(74) Representative: **Mackett, Margaret Dawn et al**  
**Rank Xerox Ltd**  
**Patent Department**  
**Parkway**  
**Marlow Buckinghamshire SL7 1YL (GB)**

(72) Inventors:  
• Bollman, James E.  
Williamson NY 14589 (US)

(54) **Automatic image cropping**

(57) The present invention describes a method for automatic cropping of images containing regions where intensity levels are uniform and other regions where intensity levels vary considerably. An image to be automatically cropped is scaled down to a grid and divided into non-overlapping blocks. The mean and variance of

an intensity level are calculated for each block. Based on the distribution of variances in the blocks, a threshold is selected for the variance. All blocks with a variance higher than this threshold variance are selected as regions of interest. The regions of interest are then cropped to a bounding rectangle.



**FIG.3A**

**EP 0 824 246 A2**

**BEST AVAILABLE COPY**

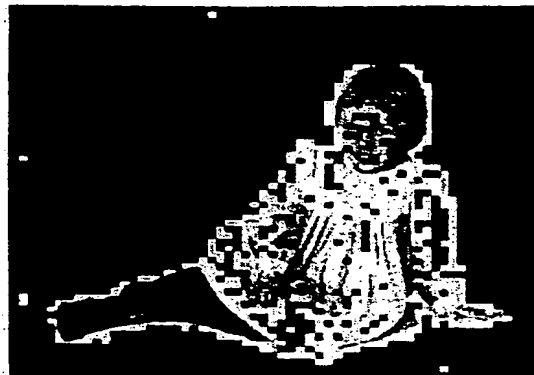


FIG. 3B



FIG. 3C



FIG. 3D

## Description

This invention is related to a method for the automatic cropping of images, and is particularly suitable to images that are texture-free, or relatively texture-free.

A typical image contains some regions where intensity level, and color, are uniform, and other regions where intensity level and color vary considerably. For instance, the "background" of an image may be uniform with a distinct "edge" separating the background from the "foreground." For example, a portrait typically comprises a subject set against a uniform backdrop or background, such that a sharp edge or boundary exists between the subject and the background.

Frequently, it is desirable to select only a particular region of an image, and to reproduce the selected region, thereby eliminating unwanted or excess background to give the image a more desirable composition. This selection process is referred to as cropping. Often, images are cropped to the foreground and most of the background is discarded.

Cropping is usually done by hand or requires operator interaction in order to properly select the subject and cropping dimensions. For example, US-A-4 809 064 discloses an apparatus for printing a selected portion of a photographic negative onto a photosensitive paper to form an enlarged and cropped photographic print. However, the apparatus requires human operation to determine the crop. Similarly, US-A-5 115 271 discloses a variable photographic cropping device for maintaining multiple constant proportions of a visible area that includes a pair of masks situated in a housing having a central window. The apparatus also requires an operator.

In the field of automatic image enhancement, methods are known for improving the contrast in a natural scene image or altering the sharpness in a reproduction of an electronically encoded natural scene images. Such methods have been disclosed, for example, in US-A-5 450 502 and US-A-5 363 209. However, such automatic image enhancement methods do not disclose automatic image cropping.

For high quality publication and printing, manual cropping may be preferred for artistic reasons. For large volume printing, including, but not limited to, passport photographs, yearbooks, catalogs, event books, portraits, other images with uniform backgrounds, and the like, it is desirable to have the option to use autocropping to enhance productivity and uniformity of the cropping process.

The present invention relates to a method for the automatic cropping of images. Further, the present invention relates to a method for automatically cropping images that are texture-free, or relatively texture-free, to their regions of interest.

In accordance with one aspect of the present invention, there is provided a method for automatically cropping an image, comprising: scaling down said image to a grid having four sides; dividing said grid into a plurality of non-overlapping blocks, calculating a mean intensity level for each of said blocks, calculating a variance of an intensity level for each of said blocks, creating a variance profile for said blocks, calculating a threshold variance based on said variance profile, selecting said blocks having said variance higher than said threshold variance as regions of interest, and cropping said regions of interest to a bounding rectangle.

In accordance with another aspect of the present invention, there is provided an apparatus for automatically cropping an image, comprising: means for scaling down said image to a grid having four sides; means for dividing said grid into a plurality of non-overlapping blocks, means for calculating a mean intensity level for each of said blocks, means for calculating a variance of an intensity level for each of said blocks, means for creating a variance profile for said blocks, means for calculating a threshold variance based on said variance profile, means for selecting said blocks having said variance higher than said threshold variance as regions of interest, and means for cropping said regions of interest to a bounding rectangle.

According to the present invention, an image to be automatically cropped is scaled down to a grid and divided into non-overlapping blocks. The mean and variance of an intensity level are calculated for each block. Based on the distribution of variances in the blocks, a threshold is selected for the variance. All blocks with a variance higher than the threshold variance are selected as regions of interest. The regions of interest are cropped to a bounding rectangle to provide an autocropped image with a tight fit. The autocropped image may then be subjected to a post-processing image operation including, but not limited to, scaling the autocropped image to a larger or smaller dimension, image enhancement, annotating, transmitting, halftoning, and the like.

The present invention may optionally include an edge strength distribution analysis of an image. A threshold is chosen from a list of edge strengths in order to select any block that contains a significant number of edge pixels and was not selected after the intensity variance analysis.

For a better understanding of the present invention, reference will now be made, by way of example only, to the accompanying drawings, in which:

FIG. 1 shows a typical intensity variance profile of an image;

FIG. 2 shows a typical edge strength profile of an image;

FIG. 3A shows a picture to be autocropped using the method of the present invention;  
 FIG. 3B shows the image after blocks with a luminance variance higher than a threshold variance are selected;  
 FIG. 3C shows the image after a post-processing cleanup pass is conducted;  
 FIG. 3D shows the autocropped image with a border scaled to the same horizontal dimension as the original picture;  
 FIG. 4A shows a second picture to be autocropped using the method of the present invention;  
 FIG. 4B shows the image after blocks with a luminance variance higher than a threshold variance are selected;  
 FIG. 4C shows the image after a post-processing cleanup pass is conducted; and  
 FIG. 4D shows the autocropped image with a border scaled to the same vertical dimension as the original picture.

The present invention relates to a method of automatic image cropping of images. Preferably, the present invention, in some embodiments, relates to automatic image cropping of images that are texture-free or relatively texture-free.

The automatic image cropping method of the present invention is independent of the image input or acquisition method. Any image acquisition device that converts a picture into electronic or digital data, such as a computer data file, is acceptable for using the method of the present invention. Acquisition devices for images to be automatically cropped include, but are not limited to, a digital scan-to-print system, a digital camera, a digital scanner, a photo CD, or laser disc, or the like.

The images themselves are defined in terms of pixels, wherein each pixel is an electrical or electronic signal with a digital gray value that varies between a white level and a black level. In a currently desirable system, in which calculations may be done on 8 bits of information or more, 256 levels of gray will be available for use. Pixels are also identified in terms of position. A pixel may define a unique location  $m,n$  within an image, identified by its  $m$ th pixel position in a line (or column), and its  $n$ th line position in a page (or row). Color is therefore represented by those gray values for red, blue and green. For example, in the RGB color space, a single color pixel is represented by three values, one for red, one for blue, and one for green. The color of the particular pixel can be defined as a combination of the red, blue and green color levels.

The automatic image cropping method of the present invention is also independent of the output method. The output methods for an image autocropped according to the present invention include, but are not limited to, a laser printer, ink jet printer, LCD display, CRT display, magnetic tape or other media, dye sublimation printer, a photographic printer, or the like. These output devices may have many characteristics. However, they have as a common requirement the representation of gray or color pictorial images.

According to an embodiment of the present invention, images are initially defined in terms of the red, green, blue ("RGB") color space. Preferably, images defined in the RGB color space are directed to a color space converter and converted to a luminance color space. In some embodiments, however, it is possible that the image will already be in luminance color space, as it is common to convert RGB values to luminance/chrominance space for other image processing. Whatever space is used, it must have a component that relates to the human visual perception of lightness or darkness.

In embodiments, the initial image data in RGB color space may be converted to a luminance color space using the luminance channel Y in Xerox YES color space of the "Xerox Color Encoding Standard", XNSS 289005, 1989. The RGB color space may also be converted to a luminance color space using the luminance channel Y in the known television encoding standard, YIQ. In an embodiment of the present invention, the luminance channel Y is calculated according to the following general formula:

$$Y = 0.252939 \times (\text{Red Channel}) + 0.684458 \times (\text{Green Channel}) + 0.062603 \times (\text{Blue Channel}).$$

All statistical analysis and color conversion is preferably carried out on a scaled down version of the input image. This speeds up the automatic image cropping process and also provides a certain degree of robustness against noise.

According to embodiments of the present invention, an image to be autocropped is scaled down to a regular grid. The grid is then divided into non-overlapping square blocks,  $N \times N$ , of a smaller size, including, but not limited to,  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$ ,  $64 \times 64$  pixels, or the like. The height and width of each block is indicated by  $N$  pixels. In an embodiment of the invention, the grid size is  $256 \times 256$  pixels, and the block size is  $4 \times 4$  pixels. However, different grid and block sizes can be used so long as the objectives of the present invention are achieved.

According to the present invention, the cue for selecting regions of interest in an image is the luminance profile of the intensity levels in the blocks. Alternatively, the RGB profile of the intensity levels in the blocks can be used. The mean and variance of the intensity level are calculated for each block consisting of  $N \times N$  pixels.

The mean intensity level,  $\mu$ , in each block is calculated according to the following formula (1):

$$\mu = \frac{1}{N^2} \sum_{i=1}^{N^2} g_i \quad (1)$$

wherein  $g_i$  is the intensity level of the  $i$ th pixel in the block, and  $N$  is the size of each block in pixels. The variance,  $\sigma$ , in each block is calculated according to the following formula (2):

$$\sigma = \frac{1}{N^2 - 1} \sum_{i=1}^{N^2} (g_i - \mu)^2 \quad (2)$$

Figure 1 shows the typical intensity variance profile of an image. From Figure 1, it is apparent that most blocks in an image exhibit very low variances.

Based on statistical analysis and the distribution of the variances in the blocks, a threshold variance is selected. From Figure 1, a generally optimal threshold variance is picked at the "knee" of the variance profile curve. According to the present invention, the threshold variance is preferably picked as a point on the curve furthest from a line joining a minimum and a maximum variance. All blocks with a variance higher than this threshold variance are selected as regions of interest (i.e., elements of the foreground) to remain in the autocropped image. All blocks with a variance less than the threshold are considered to be uninteresting (i.e., elements of the background) and are removed from the autocropped image.

The threshold variance may also be adjusted higher or lower to include or exclude more blocks. For example, the threshold variance value may be reduced by an empirically determined selectivity factor to include more objects of interest. In an embodiment of the present invention, the threshold variance is reduced by about forty percent to include more blocks as regions of interest.

The blocks selected as regions of interest are then cropped to a bounding rectangle by finding the first selected block along the four sides of the grid, thereby giving a tight fit. All blocks within the bounding rectangle are included in the autocropped image. The tightness of the crop is application dependent and is fully adjustable.

The cropped image may be scaled to a larger (or smaller) dimension and a border selected for the scaled autocropped image. In an embodiment of the present invention, the automatic cropping of an image is set to a default border of about 0.01 (i.e., a 1% border) of the larger dimension.

An optional cleanup post-processing pass may be carried out to mark unselected blocks that are inside selected regions (i.e., typically the "interior" of an object) for further post-processing image operations. An embodiment of the present invention uses a seed fill algorithm to accomplish this purpose. Various seed fill algorithms are known in the art including, but not limited to, that recited in Paul S. Heckbert, A Seed Fill Algorithm, Graphics Gems, 1990. Selected regions smaller than the specified parameter for the smallest foreground image effect that is to be retained as interesting are eliminated. In an embodiment of the present invention, small details corresponding to "noise" in the background are examined, in blocks and pixels. These details are removed from the autocropped image. As a result, small glitches and spots are eliminated, thereby providing a better bounding rectangle, especially at the edges of the autocropped image.

After the intensity variance analysis, the image optionally may be further analyzed using edge strength information as an additional cue. Some images display low contrast edges at the boundaries of the foreground and the background. Although the results of the intensity variance analysis described herein are satisfactory, results for some images may be improved if an edge strength distribution is analyzed.

Similar to the variance analysis, the edge strength distribution of the blocks is analyzed for a suitable threshold. A typical edge strength plot for an image is shown in Figure 2. In embodiments, the edge strength computation may be carried out using a digital Laplacian operator, such as that recited in Rafael C. Gonzalez & Richard E. Woods, Digital Image Processing, 100, 420, 453 (1992). Using the Laplacian operator, a threshold is chosen from the edge strength distribution. Any blocks with significant edge information (i.e., a specified number of pixels greater than the threshold) and that are not selected as regions of interest from the variance analysis are marked as "interesting".

According to the present invention, automatic image cropping relies on several empirically determined parameters for its performance. The following parameters described below include, but are not limited to, those parameters that can be tuned to customize automatic image cropping to a particular image set being analyzed. One skilled in the art may alter any or all of these parameters based on the a priori information available about the images to be autocropped.

Grid Size is the size of the square grid on which the scaled down version of the input image is sampled. Increasing the size of this grid makes the program more sensitive to noise. Image effects that would be too small to be significant on a grid of size 256 x 256, for instance, may be significant on larger grids. The time of computation also increases with the increased grid size. For example, with a 512 x 512 sampling grid, the time is approximately four times that with a 256 x 256 grid if all other parameters are unchanged.

Block Size is the size of the local neighborhood over which variance analysis is carried out. Block Size corresponds to the height and width of the non-overlapping blocks, and is typically measured in pixels. A large Block Size results in a coarse analysis of the image, while a small Block Size results in a finer analysis. The Block Size controls the size

of the local neighborhood that is included with an edge between the background and the foreground. The margin around the foreground is larger with a large Block Size.

Black Object Size is the size of the smallest background image effect that is retained as an uninteresting effect. This size is measured in blocks. All uninteresting regions whose size is less than the selected Block Size are marked as interesting regions. Typically these are completely covered by interesting regions implying that they are the interior of foreground effects.

White Object Size is the size of the smallest foreground image effect that is retained as an interesting effect. The default sizes for foreground and background effects are generally not identical, but can be independently selected. Increasing the value of this parameter has the effect of eliminating larger and larger connected regions of interest, and decreasing it has the opposite effect.

Additional parameters for an edge strength distribution analysis include, but are not limited to:

Filter Coefficients contains the filter coefficients to be used in the optional edge detection part of the automatic image cropping. Filter coefficient are weights used in a sliding window (including, but not limited to, an odd number of rows and columns) used to compute the weighted average of the grayscale values in the neighborhood of each pixel in an image. This operation can be accomplished through any edge detection filter and is not constrained to be Laplacian-like. This operation is called digital convolution and is known in the art of signal and image processing, for example in Rafael C. Gonzalez & Richard E. Woods, Digital Image Processing, (1992) and in J. Canny, IEEE Transactions in Pattern Analysis and Machine Intelligence, Vol. 8, pp. 697-98 (1986).

Edge Pixels is the number of pixels within a block that must have an edge strength greater than the threshold for the block to be selected as an interesting region. This parameter must not be disproportionately large when compared to the block size. Too small a value for this parameter makes it overly sensitive to noise. Conversely, the edge analysis phase does not add anything useful if the parameter has too high a value.

#### Examples

Two images autocropped according to the method of present invention, i.e., automatic image cropping, appear in Figures 3 and 4.

The parameters used for autocropping these images based on luminance intensity data are a Grid Size of 256 x 256 pixels; a Block Size of 4 x 4 pixels; a Black Object Size of 15 blocks; and a White Object Size of 4 blocks.

#### Example I:

Figure 3 shows an image that has been subject to the automatic image cropping method of the present invention. Figure 3A shows the original picture to be autocropped. Figure 3B shows the picture after it is subsampled into a grid (256 x 256 pixels), divided into non-overlapping blocks (64 x 64 of 4 x 4 pixels per block), and scaled to match the original. All blocks with a luminance variance higher than a threshold variance are selected as regions of interest. Figure 3C shows the autocropped image in which a post-processing pass is made to toggle groups of blocks in large areas of the opposite type of blocks and to eliminate noise in the background. Also shown is a bounding box (i.e., within dotted lines) that is calculated to give a tight fit to the autocropped image. Figure 3D shows the autocropped image, the image within the bounding box of Figure 3C, scaled to the same horizontal dimension as the original and with a 5% border.

#### Example II:

Figure 4 shows a second image that has been subject to the automatic image cropping method of the present invention. Figure 4A shows the original picture to be autocropped. Figure 4B shows the picture after it is subsampled into a grid (256 x 256 pixels), divided into non-overlapping blocks (64 x 64 of 4 x 4 pixels per block), and scaled to match the original. All blocks with a luminance variance higher than a threshold variance are selected as regions of interest. Figure 4C shows the autocropped image in which a post-processing pass is made to toggle groups of blocks in large areas of the opposite type of blocks and to eliminate noise in the background. Also shown is a bounding box (i.e., within dotted lines) that is calculated to give an autocropped image with a tight fit. Figure 4D shows the autocropped image, the image within the bonding box of Figure 4C, scaled to the same vertical dimension as the original with the side and bottom borders corresponding to the original picture and with a 5% border on the top.

Thus, automatic image cropping according to the present invention handles most texture-free, or relatively texture-free, images in a predictable and productive fashion.

It will no doubt be appreciated that the present invention can be accomplished through application software accomplishing the functions described, to operate a digital computer or microprocessor, through a hardware circuit.

The invention has been described with references to particular embodiments. Modifications and alterations will be

apparent to those skilled in the art upon reading and understanding this specification. It is intended that all such modifications and alterations are included insofar as they come within the scope of the appended claims.

## 5 Claims

1. A method for automatically cropping an image, comprising:

scaling down said image to a grid having four sides;  
dividing said grid into a plurality of non-overlapping blocks,  
calculating a mean intensity level for each of said blocks,  
calculating a variance of an intensity level for each of said blocks,  
creating a variance profile for said blocks,  
calculating a threshold variance based on said variance profile,  
selecting said blocks having said variance higher than said threshold variance as regions of interest, and  
cropping said regions of interest to a bounding rectangle.

2. The automatic image cropping method of claim 1, wherein said mean intensity level,  $\mu$ , is calculated by the formula

$$\mu = \frac{1}{N^2} \sum_{i=1}^{N^2} g_i$$

wherein  $g_i$  is the intensity level of the  $i$ th pixel in the block, and  $N$  is a height and a width of each block in pixels.

3. The automatic image cropping method of claim 2, wherein said variance,  $\sigma$ , is calculated by the formula

$$\sigma = \frac{1}{N^2 - 1} \sum_{i=1}^{N^2} (g_i - \mu)^2$$

4. The automatic image cropping method of claim 1, wherein said threshold variance is calculated as a point on a variance profile curve furthest from a line joining a minimum and a maximum variance.

5. The automatic image cropping method of claim 4, wherein said threshold variance is adjusted by a selectivity factor.

6. The automatic image cropping method of any one of claims 1 to 5, further comprising marking unselected blocks inside selected regions for further post-processing image operations.

7. The automatic image cropping method of claim 1, further comprising removing details corresponding to noise in the background of the cropped image to provide a better bounding rectangle.

8. The automatic image cropping method of any one of claims 1 to 7, further comprising:

calculating a profile of edge strengths for said blocks,  
calculating a threshold edge strength from said profile, and  
selecting said blocks having an edge strength higher than said threshold edge strength and not selected as regions of interest.

9. An apparatus for automatically cropping an image, comprising:

means for scaling down said image to a grid having four sides;  
means for dividing said grid into a plurality of non-overlapping blocks,  
means for calculating a mean intensity level for each of said blocks,  
means for calculating a variance of an intensity level for each of said blocks,  
means for creating a variance profile for said blocks,

means for calculating a threshold variance based on said variance profile,  
means for selecting said blocks having said variance higher than said threshold variance as regions of interest,  
and  
means for cropping said regions of interest to a bounding rectangle.

5

10. The apparatus of claim 9, further comprising an input means for acquiring said image and an output means for storing said autocropped image.

10

15

20

25

30

35

40

45

50

55



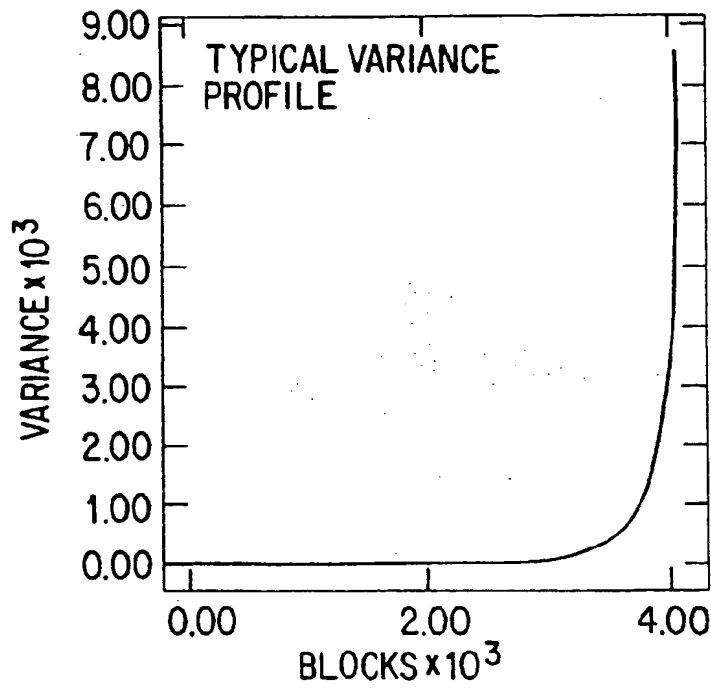


FIG. 1

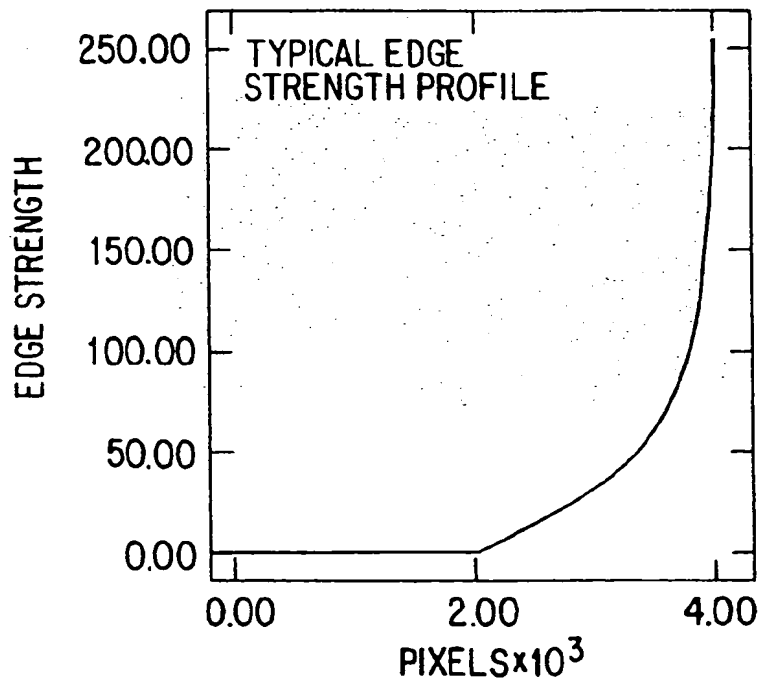


FIG. 2



FIG. 3A

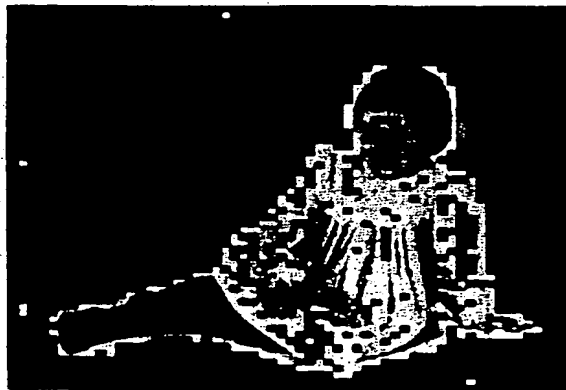


FIG. 3B

BEST AVAILABLE COPY

BEST AVAILABLE COPY



FIG. 3C



FIG. 3D

BEST AVAILABLE COPY



FIG. 4A



FIG. 4B

BEST AVAILABLE COPY



FIG. 4C



FIG. 4D